

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2003年 1月20日

出 願 番 号

Application Number:

特願2003-011591

[ST.10/C]:

[JP 2003-011591]

出 願 人

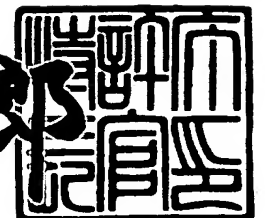
Applicant(s):

株式会社日立製作所

2003年 6月13日

特 許 庁 長 官
Commissioner,
Japan Patent Office

太田 信一郎



出証番号 出証特2003-3046495

【書類名】 特許願
【整理番号】 HI020923
【提出日】 平成15年 1月20日
【あて先】 特許庁長官殿
【国際特許分類】 G06F 3/06
【発明者】

【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社日立製作所 R A I D システム事業部内

【氏名】 杉本 守二

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 100071283

【弁理士】

【氏名又は名称】 一色 健輔

【選任した代理人】

【識別番号】 100084906

【弁理士】

【氏名又は名称】 原島 典孝

【選任した代理人】

【識別番号】 100098523

【弁理士】

【氏名又は名称】 黒川 恵

【選任した代理人】

【識別番号】 100112748

【弁理士】

【氏名又は名称】 吉田 浩二

【選任した代理人】

【識別番号】 100110009

【弁理士】

【氏名又は名称】 青木 康

【手数料の表示】

【予納台帳番号】 011785

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 記憶デバイス制御装置、及び記憶デバイス制御装置用回路基板

【特許請求の範囲】

【請求項 1】 データを記憶する記憶デバイスに接続されるとともに、ネットワークを介して情報処理装置と接続され、筐体に収納される記憶デバイス制御装置であって、

前記情報処理装置からのデータ入出力要求に対応する I/O 要求を、前記記憶デバイスに対して出力する I/O プロセッサが形成された、前記筐体に収納される回路基板を有し、

前記回路基板は、前記筐体の内方側に位置すべき端部に内方側コネクタを有し

前記内方側コネクタを通じて、少なくとも前記 I/O プロセッサと電源装置とが接続される、

記憶デバイス制御装置において、

前記回路基板には、前記データ入出力要求をファイル単位で受け付けるファイルアクセス処理部が形成され、

前記回路基板は、前記筐体の外方側に位置すべき端部に前記ファイルアクセス処理部へ電力を供給するための電力用コネクタを備えている、

ことを特徴とする記憶デバイス制御装置。

【請求項 2】 前記筐体は、前記内方側コネクタを介して前記回路基板を支持する支持部を備え、

前記支持部は、前記電源装置から少なくとも前記 I/O プロセッサに電力を供給するための導電部の一部を備えている、

ことを特徴とする請求項 1 に記載の記憶デバイス制御装置。

【請求項 3】 前記回路基板は、前記筐体の外方側に位置すべき端部に通信用コネクタを有し、前記通信用コネクタを通じて、前記ファイルアクセス処理部と前記情報処理装置とが接続されることを特徴とする請求項 1 又は 2 に記載の記憶デバイス制御装置。

【請求項 4】 前記筐体は、収納される前記回路基板に対して空気流を形成

するための送風部を備えていることを特徴とする請求項1乃至3のいずれかに記載の記憶デバイス制御装置。

【請求項5】 前記ファイルアクセス処理部は、前記回路基板における前記空気流の上流側に形成されていることを特徴とする請求項4に記載の記憶デバイス制御装置。

【請求項6】 前記ファイルアクセス処理部と前記I/Oプロセッサとは、互いに平行で重なった状態にある複数枚の回路基板を一組とする回路基板部に、互いに分けて形成されていることを特徴とする請求項4又は5に記載の記憶デバイス制御装置。

【請求項7】 前記回路基板において、前記ファイルアクセス処理部と前記電力用コネクタとの間に、前記電源装置から入力される電圧を降下させて前記ファイルアクセス処理部に供給する電圧変換部が形成されていることを特徴とする請求項1乃至6のいずれかに記載の記憶デバイス制御装置。

【請求項8】 前記電源装置から電力用ケーブルを介して、電力が前記電力用コネクタへ供給されることを特徴とする請求項1乃至7のいずれかに記載の記憶デバイス制御装置。

【請求項9】 データを記憶する記憶デバイスに接続されるとともに、ネットワークを介して情報処理装置と接続され、筐体に収納される記憶デバイス制御装置に用いられる回路基板であって、

前記情報処理装置からのデータ入出力要求に対応するI/O要求を、前記記憶デバイスに対して出力するI/Oプロセッサが形成され、

前記筐体に収納されるとともに、

前記筐体の内方側に位置すべき端部に内方側コネクタを有し、

前記内方側コネクタを通じて、少なくとも前記I/Oプロセッサと電源装置とが接続される、

回路基板において、

前記データ入出力要求をファイル単位で受け付けるファイルアクセス処理部が形成され、

前記筐体の外方側に位置すべき端部に前記ファイルアクセス処理部へ電力を供

給するための電力用コネクタを備えている、

ことを特徴とする記憶デバイス制御装置用回路基板。

【請求項10】 前記筐体は、前記内方側コネクタを介して前記回路基板を支持する支持部を備え、

前記支持部は、前記電源装置から少なくとも前記I/Oプロセッサに電力を供給するための導電部の一部を備えている、

ことを特徴とする請求項9に記載の記憶デバイス制御装置用回路基板。

【請求項11】 前記筐体の外方側に位置すべき端部に通信用コネクタを有し、前記通信用コネクタを通じて、前記ファイルアクセス処理部と前記情報処理装置とが接続されることを特徴とする請求項9又は10に記載の記憶デバイス制御装置用回路基板。

【請求項12】 前記筐体は、収納される前記回路基板に対して空気流を形成するための送風部を備えていることを特徴とする請求項9乃至11のいずれかに記載の記憶デバイス制御装置用回路基板。

【請求項13】 前記ファイルアクセス処理部は、前記空気流の上流側に形成されていることを特徴とする請求項12に記載の記憶デバイス制御装置用回路基板。

【請求項14】 互いに平行で重なった状態にある複数枚を一組としてなる回路基板部に対して、前記ファイルアクセス処理部と前記I/Oプロセッサとは、互いに分けて形成されていることを特徴とする請求項12又は13に記載の記憶デバイス制御装置用回路基板。

【請求項15】 前記ファイルアクセス処理部と前記電力用コネクタとの間に、前記電源装置から出力される電圧を降下させて前記ファイルアクセス処理部に印加する電圧変換部が形成されていることを特徴とする請求項9乃至14のいずれかに記載の記憶デバイス制御装置用回路基板。

【請求項16】 前記電源装置から電力用ケーブルを介して、電力が前記電力用コネクタへ供給されることを特徴とする請求項9乃至15のいずれかに記載の記憶デバイス制御装置用回路基板。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、記憶デバイス制御装置、及び記憶デバイス制御装置用回路基板に関する。

【0002】

【従来の技術】

近年、コンピュータシステムで取り扱われるデータ量が急激に増加している。かかる膨大なデータを効率よく利用し管理するために、複数のディスクアレイ装置（以後、ストレージシステムと称する）と情報処理装置とを専用のネットワーク（Storage Area Network、以後SANと記す）で接続し、ストレージシステムへの高速かつ大量なアクセスを実現する技術が開発されている。ストレージシステムと情報処理装置とをSANで接続し、高速なデータ転送を実現するためには、ファイバチャネルプロトコルに従った通信機器を用いてネットワークを構築するのが一般的である。

【0003】

一方、複数のストレージシステムと情報処理装置とを、TCP/IP (Transmission Control Protocol/Internet Protocol) プロトコルを用いたネットワークで相互に接続し、ストレージシステムへのファイルレベルでのアクセスを実現する、NAS (Network Attached Storage) と呼ばれるネットワークシステムが開発されている。NASにおいては、ストレージシステムに対してファイルシステム機能を有する装置が接続されているため、情報処理装置からのファイルレベルでのアクセスが可能となっている。特に最近では、ミッドレンジクラスやエンタープライズクラスと呼ばれるような、巨大な記憶資源を提供するRAID (Redundant Arrays of Inexpensive Disks) 方式で管理されたストレージシステムにファイルシステムを結合させた、大規模なNASが注目されている。

【0004】

【特許文献1】

特開2002-351703号公報

【0005】

【発明が解決しようとする課題】

しかしながら、従来のNASは、TCP/IP通信機能及びファイルシステム機能を持たないストレージシステムに、TCP/IP通信機能及びファイルシステム機能を持った情報処理装置を接続させることにより実現されていた。このため、接続される情報処理装置の設置スペースが必要であった。従って、このような設置スペース分だけNASの占める床面積が大きくなるために、NASの設置面積当たりのデータ記憶容量が相対的に低下する恐れがあった。また、従来のNASにおいては、前記情報処理装置とストレージシステムとの間は、高速に通信を行なう必要性から、SANで接続されていることが多く、このための通信制御機器や通信制御機能を備える必要もあった。

【0006】

NASの設置面積当たりのデータ記憶容量を向上させ、SANに関連する機器や機能の煩雑さを低減させるために、従来のNASにおける従来の情報処理装置とのインタフェース機能が搭載された回路基板に対してTCP/IP機能やファイルシステム機能を実現する手段をさらに搭載すると仮定する。この場合、この回路基板は、情報処理装置やストレージシステム等に対して新たな接続態様を有する回路基板としなければならない。従って、従来の回路基板が収納されていた筐体も、前記の新しい接続態様に対応する筐体としなければならない。これによりNASは高コストとなり、また、回路基板及び筐体をともに新規にすることはユーザへの大きな負担となる。

【0007】

本発明はかかる課題に鑑みてなされたものであり、その目的とするところは、記憶デバイス制御装置、及び記憶デバイス制御装置用回路基板を提供することにある。

【0008】

【課題を解決するための手段】

本発明のうち主たる発明は、データを記憶する記憶デバイスに接続されるとともに、ネットワークを介して情報処理装置と接続され、筐体に収納される記憶デバイス制御装置であって、前記情報処理装置からのデータ入出力要求に対応する

I/O要求を、前記記憶デバイスに対して出力するI/Oプロセッサが形成された、前記筐体に収納される回路基板を有し、前記回路基板は、前記筐体の内方側に位置すべき端部に内方側コネクタを有し、前記内方側コネクタを通じて、少なくとも前記I/Oプロセッサと電源装置とが接続される、記憶デバイス制御装置において、前記回路基板には、前記データ入出力要求をファイル単位で受け付けるファイルアクセス処理部が形成され、前記回路基板は、前記筐体の外方側に位置すべき端部に前記ファイルアクセス処理部へ電力を供給するための電力用コネクタを備えている。

【0009】

ここで、前記情報処理装置からのデータ入出力要求に対応するI/O要求を、前記記憶デバイスに対して出力するI/Oプロセッサが形成された、前記筐体に収納される回路基板であって、且つ前記データ入出力要求をファイル単位で受け付けるファイルアクセス処理部が形成される回路基板とは、例えばNASにおける後述するチャンネル制御部の回路基板、即ちNASボードのことである。また、ファイルアクセス処理部は、NFS (Network File System) やSamba等のファイルシステムプロトコルを含んで構成されるネットワーク制御部を少なくとも含む。NFSは、UNIX (登録商標) 系の情報処理装置からのファイルアクセス要求を受け付けるためのファイルシステムプロトコルである。また、SambaはCIFS (Common Interface File System) が動作するWindows (登録商標) 系の情報処理装置からのファイルアクセス要求を受け付けるためのファイルシステムプロトコルである。尚、ファイルアクセス処理部には、その他のソフトウェアが含まれるようにすることもできる。

【0010】

前記筐体の外方側に位置すべき端部に前記ファイルアクセス処理部へ電力を供給するための電力用コネクタを前記回路基板が備えているということは、例えばチャンネル制御部の回路基板に形成されたファイルアクセス処理部へ給電するためのコネクタが、例えば回路基板の筐体正面側に取り付けられているということである。

【0011】

前記の発明によれば、例えば従来はそれぞれ別筐体にあった、ファイルアクセス処理部による、データ入出力要求をファイル単位で受け付ける機能と、I/Oプロセッサによる、I/O要求を出力する機能とを一体をなす回路基板で実現することによって、NAS全体がコンパクトになる。

【0012】

また、ファイルアクセス処理部への給電を、回路基板の内方側コネクタとは別に、電力用コネクタから行なうことによって、従来のNASの回路基板を大きく改造することなく、前述の2つの機能を実現する回路基板とすることができる。何故なら、従来のNASの回路基板に取り付けられている既存のコネクタにおける給電のための箇所をほとんど改造せずに、本発明のNASの回路基板にそのまま使用できるからである。従って、従来の回路基板に新たにファイルアクセス処理部を形成し、本発明のNASにおけるチャネル制御部として使用できる。同様に、従来のNASにおける既存の筐体も大きく改造せずに、本発明によるNASのチャネル制御部の回路基板を装着して使用できる。

【0013】

また、本発明のうち主たる発明は、データを記憶する記憶デバイスに接続されるとともに、ネットワークを介して情報処理装置と接続され、筐体に収納される記憶デバイス制御装置に用いられる回路基板であって、前記情報処理装置からのデータ入出力要求に対応するI/O要求を、前記記憶デバイスに対して出力するI/Oプロセッサが形成され、前記筐体に収納されるとともに、前記筐体の内方側に位置すべき端部に内方側コネクタを有し、前記内方側コネクタを通じて、少なくとも前記I/Oプロセッサと電源装置とが接続される、回路基板において、前記データ入出力要求をファイル単位で受け付けるファイルアクセス処理部が形成され、前記筐体の外方側に位置すべき端部に前記ファイルアクセス処理部へ電力を供給するための電力用コネクタを備えている。

【0014】

前記記憶デバイス制御装置用回路基板とは、例えばNASにおけるチャネル制御部の回路基板、即ちNASボードのことである。

【0015】

【発明の実施の形態】

===ストレージシステムの概要===

<<<全体構成例>>>

本発明の実施の形態について、図面を用いて詳細に説明する。

図1において、本実施の形態における記憶デバイス制御装置を備えるストレージシステムの全体構成を示す。

本実施の形態において、ストレージシステム600は、記憶デバイス制御装置100と記憶デバイス300とを備えている。記憶デバイス制御装置100は、情報処理装置200から受信したコマンドに従って記憶デバイス300に対する制御を行なう。例えば、情報処理装置200からデータの入出力要求を受信して、記憶デバイス300に記憶されているデータの入出力のための処理を行なう。データは、記憶デバイス300が備えるディスクドライブにより提供される物理的な記憶領域上に論理的に設定される記憶領域である論理ボリューム(Logical Unit) (以後、LUと記す) に記憶されている。また、記憶デバイス制御装置100は、情報処理装置200との間で、ストレージシステム600を管理するための各種コマンドの授受も行なう。

【0016】

情報処理装置200は、CPU (Central Processing Unit) やメモリを備えたコンピュータである。情報処理装置200が備えるCPUによって各種プログラムが実行され、これにより様々な機能が実現される。情報処理装置200は、例えばパーソナルコンピュータやワークステーションとすることもできるし、メインフレームコンピュータとすることもできる。

【0017】

図1において、情報処理装置1乃至3 (200) は、LAN (Local Area Network) 400を介して記憶デバイス制御装置100と接続されている。LAN 400は、インターネットとすることもできるし、専用のネットワークとすることもできる。LAN 400を介して行われる情報処理装置1乃至3 (200) と記憶デバイス制御装置100との間の通信は、例えばTCP/IPプロトコルに従って行われる。情報処理装置1乃至3 (200) からは、ストレージシステム6

00に対して、ファイル名指定によるデータアクセス要求（ファイル単位でのデータ入出力要求。以後、ファイルアクセス要求と記す）が送信される。

【0018】

LAN400には、バックアップデバイス910が接続されている。バックアップデバイス910は、具体的にはMOやCD-R、DVD-RAM等のディスク系デバイス、DATテープ、カセットテープ、オープンテープ、カートリッジテープ等のテープ系デバイスである。バックアップデバイス910は、LAN400を介して記憶デバイス制御装置100との間で通信を行なうことにより、記憶デバイス300に記憶されているデータのバックアップデータを記憶する。また、テープデバイス910は、情報処理装置1（200）と接続されるようにすることもできる。この場合は、情報処理装置1（200）を介して、記憶デバイス300に記憶されているデータのバックアップデータを取得するようにする。

【0019】

また、記憶デバイス制御装置100は、チャンネル制御部1乃至4（110）を備える。記憶デバイス制御装置100は、チャンネル制御部1乃至4（110）によりLAN400を介して情報処理装置1乃至3（200）やバックアップデバイス910との間で通信を行なう。チャンネル制御部1乃至4（110）は、情報処理装置1乃至3（200）からのファイルアクセス要求を個々に受け付ける。即ち、チャンネル制御部1乃至4（110）には、個々にLAN400上のネットワークアドレス（例えば、IPアドレス）が割り当てられていて、それぞれが個別にNASとして振る舞い、個々のNASが、あたかも独立したNASが存在するかのように、NASとしてのサービスを情報処理装置1乃至3（200）に提供することができる。以後、チャンネル制御部1乃至4（110）をCHNと記す。このように、1台のストレージシステム600に、個別にNASとしてのサービスを提供するCHN1乃至4（110）を備えるように構成したことで、従来、独立したコンピュータで個々に運用されていたNASサーバが1台のストレージシステム600に集約される。そして、これにより、ストレージシステム600の統括的な管理が可能となり、各種設定・制御や障害管理、バージョン管理といった保守業務の効率化が図られる。

【0020】

尚、本実施の形態における記憶デバイス制御装置100のCHN1乃至4（110）は、後述するように、一体的にユニット化された回路基板上に形成されたハードウェア、及びこのハードウェアにより実行される、オペレーティングシステム（以後、OSと記す）や、このOS上で動作するアプリケーションプログラム等のソフトウェアにより実現される。このように、本実施の形態のストレージシステム600では、従来ハードウェアの一部として実装されてきた機能が、ソフトウェアにより実現されている。このため、本実施の形態の記憶制御装置システム600では、柔軟性に富んだシステム運用が可能となり、多様で変化の激しいユーザニーズに対し、よりきめ細かなサービスの提供が可能となる。

【0021】

情報処理装置3及び4（200）は、SAN（Storage Area Network）500を介して記憶デバイス制御装置100と接続されている。SAN500は、記憶デバイス300が提供する記憶領域におけるデータの管理単位であるブロックを単位として情報処理装置3及び4（200）との間でデータの授受を行なうためのネットワークである。SAN500を介して行われる情報処理装置3及び4（200）と記憶デバイス制御装置100との間の通信は、一般にファイバチャネルプロトコルに従って行われる。情報処理装置3及び4（200）からは、ストレージシステム600に対して、ファイバチャネルプロトコルに従ってブロック単位のデータアクセス要求（以後、ブロックアクセス要求と記す）が送信される。

【0022】

SAN500には、SAN対応のバックアップデバイス900が接続されている。SAN対応バックアップデバイス900は、SAN500を介して記憶デバイス制御装置100との間で通信を行なうことにより、記憶デバイス300に記憶されているデータのバックアップデータを記憶する。

【0023】

記憶デバイス制御装置100は、チャネル制御部5及び6（110）を備える。記憶デバイス制御装置100は、チャネル制御部5及び6（110）によりS

AN500を介して情報処理装置3及び4(200)及びSAN対応バックアップデバイス900との間の通信を行なう。以後、チャンネル制御部5及び6(110)をCHFと記す。

【0024】

また、情報処理装置5(200)は、LAN400やSAN500等のネットワークを介さずに記憶デバイス制御装置100と接続されている。情報処理装置5(200)としては、例えば、メインフレームコンピュータとすることができる。情報処理装置5(200)と記憶デバイス制御装置100との間の通信は、例えば、FICON(Fibre Connection)(登録商標)やESCON(Enterprise System Connection)(登録商標)、ACONARC(Advanced Connection Architecture)(登録商標)、FIBARC(Fibre Connection Architecture)(登録商標)等の通信プロトコルに従って行われる。情報処理装置5(200)からは、ストレージシステム600に対して、これらの通信プロトコルに従ってブロックアクセス要求が送信される。

【0025】

記憶デバイス制御装置100は、チャンネル制御部7及び8(110)により情報処理装置5(200)との間で通信を行なう。以後、チャンネル制御部7及び8(110)をCHAと記す。

【0026】

SAN500には、ストレージシステム600の設置場所(プライマリサイト)とは遠隔した場所(セカンダリサイト)に設置される他のストレージシステム610が接続している。ストレージシステム610は、後述するレプリケーション又はリモートコピーの機能におけるデータの複製先の装置として利用される。尚、ストレージシステム610は、SAN500以外にも、ATM等の通信回線によりストレージシステム600に接続していることもある。この場合には、例えば、チャンネル制御部110として、前記通信回線を利用するためのインタフェース(チャンネルエクステンダ)を備えるチャンネル制御部110が採用される。

【0027】

本実施の形態によれば、ストレージシステム600内にCHN110、CHF

110、CHA110を混在させて装着させることにより、異種ネットワークに接続されるストレージシステムを実現できる。具体的には、ストレージシステム600は、CHN110を用いてLAN140に接続し、且つ、CHF110を用いてSAN500に接続するという、SAN-NAS統合ストレージシステムである。

【0028】

<<<記憶デバイス>>>

図1における記憶デバイス300は、多数のディスクドライブ（物理ディスク）を備えており、情報処理装置200に対して記憶領域を提供する。データは、ディスクドライブにより提供される物理的な記憶領域上に論理的に設定される記憶領域であるLUに記憶されている。ディスクドライブとしては、例えば、ハードディスク装置やフレキシブルディスク装置、半導体記憶装置等、様々なものを用いることができる。尚、記憶デバイス300は、例えば複数のディスクドライブによりディスクアレイを構成するようにすることもできる。この場合、情報処理装置200に対して提供される記憶領域は、RAIDにより管理された複数のディスクドライブにより提供されるようにすることもできる。

【0029】

また、記憶デバイス制御装置100と記憶デバイス300との間は、図1のように直接に接続される形態とすることもできるし、ネットワークを介して接続するようにすることもできる。さらに記憶デバイス300を、記憶デバイス制御装置100と一体型として構成することもできる。

【0030】

記憶デバイス300に設定されるLUには、情報処理装置200からアクセス可能なユーザLUや、チャネル制御部110の制御のために使用されるシステムLU等がある。システムLUには、チャネル制御部110で実行されるオペレーティングシステムも格納される。また、各LUには、チャネル制御部110が対応付けられている。これにより、チャネル制御部110毎にアクセス可能なLUが割り当てられている。また前記対応付けは、複数のチャネル制御部110で1つのLUを共有するようにすることもできる。尚、以後、ユーザLUやシステム

LUをユーザディスク、システムディスク等とも記す。また、複数のチャネル制御部110で共有されるLUを、共有LU或いは共有ディスクと記す。

【0031】

<<<記憶デバイス制御装置>>>

図1に示されるように、記憶デバイス制御装置100は、チャネル制御部110、共有メモリ120、キャッシュメモリ130、ディスク制御部140、管理端末160、及び接続部150を備える。

【0032】

チャネル制御部110は、情報処理装置200との間で通信を行なうための通信インタフェースを備え、情報処理装置200との間でデータ入出力コマンド等を授受する機能を有する。例えば、CHN110は、情報処理装置1乃至3(200)からのファイルアクセス要求を受け付ける。そして、ファイルの記憶アドレスやデータ長等を求めて、ファイルアクセス要求に対応するI/O要求を出力することにより、記憶デバイスへのアクセスを行なう。これによりストレージシステム600は、NASとしてのサービスを情報処理装置1乃至3(200)に提供することができる。尚、I/O要求には、データの先頭のアドレス、データ長、読み出し又は書き込み等のアクセスの種別が含まれている。また、データの書き込みの場合には、I/O要求には書き込みデータが含まれているようにすることもできる。I/O要求の出力は、後述するI/Oプロセッサ119により行なわれる。また、CHF110は、情報処理装置3及び4(200)からのファイバチャネルプロトコルに従ったブロックアクセス要求を受け付ける。これにより、ストレージシステム600は、高速アクセス可能なデータ記憶サービスを情報処理装置3及び4(200)に対して提供することができる。またCHA110は、情報処理装置5(200)からのFICONやESCON、ACONARC、FIBARC等のプロトコルに従ったブロックアクセス要求を受け付ける。これにより、ストレージシステム600は、情報処理装置5(200)のようなメインフレームコンピュータに対してもデータ記憶サービスを提供することができる。

【0033】

各チャンネル制御部110は、管理端末160とともに内部LAN151で接続されている。これにより、チャンネル制御部110に実行させるマイクロプログラム等を管理端末160から送信しインストールすることが可能となっている。チャンネル制御部110の構成については後述する。

【0034】

図1において、接続部150は、チャンネル制御部110、共有メモリ120、キャッシュメモリ130、及びディスク制御部140を相互に接続する。チャンネル制御部110、共有メモリ120、キャッシュメモリ130、及びディスク制御部140の間でのデータやコマンドの授受は接続部150を介することにより行われる。接続部150は、例えば、高速スイッチングによりデータ伝送を行なうクロスバススイッチ等の高速バスで構成される。チャンネル制御部110どうしが高速バスで接続されていることで、個々のコンピュータ上で動作するNASサーバが、LANを通じて接続する従来の構成に比べてチャンネル制御部110間の通信パフォーマンスが大幅に向上する。また、これにより、高速なファイル共有機能や高速フェイルオーバ等が可能となる。

【0035】

共有メモリ120及びキャッシュメモリ130は、チャンネル制御部110及びディスク制御部140により共有される記憶メモリである。共有メモリ120は、主に制御情報やコマンド等を記憶するために利用されるのに対し、キャッシュメモリ130は、主にデータを記憶するために利用される。

【0036】

例えば、あるチャンネル制御部110が情報処理装置200から受信したデータ入出力コマンドが書き込みコマンドであった場合には、当該チャンネル制御部110は書き込みコマンドを共有メモリ120に書き込むとともに、情報処理装置200から受信した書き込みデータをキャッシュメモリ130に書き込む。一方、ディスク制御部140は共有メモリ120を監視しており、共有メモリ120に書き込みコマンドが書き込まれたことを検出すると、当該コマンドに従ってキャッシュメモリ130から書き込みデータを読み出して記憶デバイス300に書き込む。

【0037】

また、あるチャネル制御部110により情報処理装置200から受信されたデータ入出力コマンドが読み出しコマンドであった場合には、当該チャネル制御部110は、読み出しコマンドを共有メモリ120に書き込むとともに、読み出し対象となるデータがキャッシュメモリ130に存在するか否かを調べる。ここで、当該データがキャッシュメモリ130に存在すれば、チャネル制御部110は、そのデータを情報処理装置200に送信する。一方、読みだし対象となるデータがキャッシュメモリ130に存在しない場合には、共有メモリ120の監視により、読み出しコマンドの共有メモリ120への書き込みを検出したディスク制御部140が、記憶デバイス300から読み出し対象となるデータを読み出してこれをキャッシュメモリ130に書き込むとともに、その旨を共有メモリ120に書き込む。そして、チャネル制御部110は、共有メモリ120を監視することにより、読み出し対象となるデータのキャッシュメモリ130への書き込みを検出すると、そのデータを情報処理装置200に送信する。

【0038】

尚、このように、チャネル制御部110からディスク制御部140に対するデータの書き込みや読み出しの指示を共有メモリ120を介在させて間接的に行なう構成の他、例えば、チャネル制御部110からディスク制御部140に対してデータの書き込みや読み出しの指示を共有メモリ120を介さずに直接的に行なう構成とすることもできる。

【0039】

ディスク制御部140は、記憶デバイス300の制御を行なう。例えば前述のように、チャネル制御部110が情報処理装置200から受信したデータ書き込みコマンドに従って記憶デバイス300へデータの書き込みを行なう。また、チャネル制御部110により送信された論理アドレス指定によるLUへのデータアクセス要求を、物理アドレス指定による物理ディスクへのデータアクセス要求に変換する。記憶デバイス300における物理ディスクがRAIDにより管理されている場合には、RAID構成に従ったデータのアクセスを行なう。また、ディスク制御部140は、記憶デバイス300に記憶されたデータの複製管理の制御

やバックアップ制御を行なう。さらに、ディスク制御部140は、災害発生時のデータ消失防止（ディザスタリカバリ）等を目的としてプライマリサイトのストレージシステム600のデータ複製をセカンダリサイトに設置された他のストレージシステム610にも記憶する制御（レプリケーション機能、又はリモートコピー機能）等も行なう。

【0040】

各ディスク制御部140は、管理端末160とともに、内部LAN151を介して相互に接続され、相互に通信を行なうことが可能である。これにより、ディスク制御部140に実行させるマイクロプログラム等を管理端末160から送信しインストールすることが可能となっている。ディスク制御部の構成については、後述する。

【0041】

本実施の形態においては、共有メモリ120及びキャッシュメモリ130が、チャンネル制御部110及びディスク制御部140に対して独立に設けられていることについて記載したが、本実施の形態はこの場合に限られるものでなく、共有メモリ120又はキャッシュメモリ130が、チャンネル制御部110及びディスク制御部140の各々に分散されて設けられることも好ましい。この場合、接続部150は、分散された共有メモリ又はキャッシュメモリを有するチャンネル制御部110及びディスク制御部140を相互に接続させることになる。

【0042】

<<<管理端末>>>

図1における管理端末160は、ストレージシステム600を保守・管理するためのコンピュータである。管理端末160を操作することにより、例えば記憶デバイス300内の物理ディスク構成の設定や、LUの設定、チャンネル制御部110において実行されるマイクロプログラムのインストール等を行なうことができる。ここで、記憶デバイス300内の物理ディスク構成の設定としては、例えば物理ディスクの増設や減設、RAID構成の変更（RAID1からRAID5への変更等）等を行なうことができる。さらに、管理端末160からは、ストレージシステム600の動作状態の確認や故障部位の特定、チャンネル制御部110

で実行されるオペレーティングシステムのインストール等の作業を行なうこともできる。また、管理端末160は、LANや電話回線等で外部保守センタと接続されており、管理端末160を利用してストレージシステム600の障害監視を行ったり、障害が発生した場合には、これに迅速に対応したりすることが可能である。障害の発生は、例えばOSやアプリケーションプログラム、ドライバソフトウェア等から通知される。この通知は、HTTPプロトコルや、SNMP (Simple Network Management Protocol)、電子メール等により行なわれる。これらの設定や制御は、管理端末160で動作するWebサーバが提供するWebページをユーザインタフェースとしてオペレータ等により行われる。オペレータ等は、管理端末160を操作して、障害監視する対象や内容の設定、障害通知先の設定等を行なうこともできる。

【0043】

管理端末160は、記憶デバイス制御装置100に内蔵されている形態とすることもできるし、外付けされている形態とすることもできる。また、管理端末160は、記憶デバイス制御装置100及び記憶デバイス300の保守・管理を専用に行なうコンピュータとすることもできるし、汎用のコンピュータに保守・管理機能を持たせたものとすることもできる。

【0044】

図2において、管理端末160の構成を示す。管理端末160は、CPU161、メモリ162、ポート163、記録媒体読取装置164、入力装置165、出力装置166、及び記憶装置168を備える。

【0045】

CPU161は、管理端末160の全体の制御を司るもので、メモリ162に格納されたプログラム162cを実行することにより前記Webサーバとしての機能等を実現する。メモリ162には、物理ディスク管理テーブル162aとLU管理テーブル162bとプログラム162cとが記憶されている。

【0046】

物理ディスク管理テーブル162aは、記憶デバイス300に備えられる物理ディスク（ディスクドライブ）を管理するためのテーブルである。

図3において、物理ディスク管理テーブル162aを示す。図3においては、記憶デバイス300が備える多数の物理ディスクのうち、ディスク番号#001乃至#006までが示されている。それぞれの物理ディスクに対して、容量、RAID構成、使用状況が示されている。

【0047】

LU管理テーブル162bは、前記物理ディスク上に論理的に設定されるLUを管理するためのテーブルである。

図4において、LU管理テーブル162bを示す。図4においては、記憶デバイス300上に設定される多数のLUのうち、LU番号#1乃至#3までが示されている。それぞれのLUに対して、物理ディスク番号、容量、RAID構成が示されている。

【0048】

記録媒体読取装置164は、記録媒体167に記録されているプログラムやデータを読み取るための装置である。読み取られたプログラムやデータは、メモリ162や記憶装置168に格納される。従って、例えば、記録媒体167に記録されたプログラム162cを、記録媒体読取装置164を用いて前記記録媒体167から読み取って、メモリ162や記憶装置168に格納するようにすることができる。記録媒体167としては、フレキシブルディスクやCD-ROM、DVD-ROM、DVD-RAM、半導体メモリ等を用いることができる。尚、前記プログラム162cは、管理端末160を動作させるためのプログラムとすることができる他、チャンネル制御部110やディスク制御部140にOS701やアプリケーションプログラムをインストールするためのプログラムや、バージョンアップするためのプログラムとすることもできる。記録媒体読取装置164は、管理端末160に内蔵されている形態とすることもできるし、外付されている形態とすることもできる。記憶装置168は、例えば、ハードディスク装置やフレキシブルディスク装置、半導体記憶装置等である。入力装置165は、オペレータ等による管理端末160へのデータ入力等のために用いられる。入力装置165としては、例えば、キーボードやマウス等が用いられる。出力装置166は、情報を外部に出力するための装置である。出力装置166としては、例えば、

ディスプレイやプリンタ等が用いられる。ポート163は、内部LAN151に接続されており、これにより、管理端末160は、チャネル制御部110やディスク制御部140等と通信を行なうことができる。またポート163は、LAN400に接続するようにすることもできるし、電話回線に接続するようにすることもできる。

【0049】

<<<ストレージシステムの外観>>>

図5において、本実施の形態における記憶デバイス制御装置100を備えるストレージシステム600の外観構成を示す。図5に示すように、ストレージシステム600は、記憶デバイス制御装置100及び記憶デバイス300がそれぞれの筐体に納められた形態をなしている。記憶デバイス制御装置100の筐体の両側に記憶デバイス300の筐体が配置されている。

【0050】

図6において、記憶デバイス制御装置100の外観構成を示す。図6において、記憶デバイス制御装置100の正面及び上面、並びに当該記憶デバイス制御装置の筐体101の正面、上面、及び一側面が示されている。本実施の形態においては、図6における筐体101の背面側には、記憶デバイス300が収納されている。尚、図6における筐体101の背面側に、記憶デバイス制御装置100を正面側と対称に収納し、記憶デバイス制御装置100を集約させてもよい。この場合、記憶デバイス制御装置100と、別筐体に収納される記憶デバイス300とは、ケーブル等で電氣的に接続される。

【0051】

記憶デバイス制御装置100には、図6における第2段目ボックス103に隣接するように、管理端末160が備えられている。図6に示した管理端末160は、いわゆるノート型パーソナルコンピュータの形態をしているが、どのような形態であってもよい。

【0052】

図6において、筐体101の第2段目ボックス103には、チャネル制御部110を装着するためのスロットが設けられている。各スロットにはチャネル制御

部 1 1 0 のボードが装着される。チャンネル制御部 1 1 0 のボードとは、チャンネル制御部 1 1 0 の回路基板が形成されたユニットであり、スロットへの装着単位である。また、各スロットにはチャンネル制御部 1 1 0 のボードを装着するためのガイドレールが設けられている。ガイドレールに沿ってチャンネル制御部 1 1 0 のボードをスロットに挿入することにより、チャンネル制御部 1 1 0 を記憶デバイス制御装置 1 0 0 の一部とすることができる。また各スロットに装着されたチャンネル制御部 1 1 0 のボードは、ガイドレールに沿って手前方向に引き抜くことにより取り外すことができる。また各スロットの奥手側（内方側）には、プラッタ（支持部）が設けられ、このプラッタには、各チャンネル制御部 1 1 0 を記憶デバイス制御装置 1 0 0 の一部として電氣的に接続するためのコネクタが設けられている。

【 0 0 5 3 】

ここで、チャンネル制御部 1 1 0 には、CHN、CHF、CHAがあるが、いずれのチャンネル制御部 1 1 0 のボードも、サイズやコネクタの位置、コネクタのピン配列等に互換性をもたせているため、8つのスロットにはいずれのチャンネル制御部 1 1 0 のボードも装着することが可能である。従って、例えば8つのスロット全てにCHN 1 1 0 のボードを装着するようにすることもできる。また図 1 に示したように、4枚のCHN 1 1 0 のボードと、2枚のCHF 1 1 0 のボードと、2枚のCHA 1 1 0 のボードとを装着するようにすることもできる。チャンネル制御部 1 1 0 のボードを装着しないスロットを設けるようにすることもできる。

【 0 0 5 4 】

本実施の形態においては、各スロットにボードが装着されたチャンネル制御部 1 1 0 は、同種の複数のチャンネル制御部 1 1 0 でクラスタを構成する。例えば2つのチャンネル制御部 1 1 0 を一対としてクラスタを構成することができる。クラスタを構成することにより、クラスタ内の、あるチャンネル制御部 1 1 0 に障害が発生した場合でも、障害が発生したチャンネル制御部 1 1 0 がそれまで行っていた処理をクラスタ内の他のチャンネル制御部 1 1 0 に引き継ぐようにすることができる（フェイルオーバー制御）。

図 1 2 において、2つのチャンネル制御部 1 1 0 をクラスタで構成している様子

を示すが、詳細は後述する。

【0055】

尚、記憶デバイス制御装置100は、信頼性向上のため、電源供給が2系統化されており、チャンネル制御部110のボードが装着される前記8つのスロットは、電源系統毎に4つずつに分けられている。ここでクラスタを構成する場合には、両方の電源系統のチャンネル制御部110のボードを含むようにする。これにより、片方の電源系統に障害が発生し電力の供給が停止しても、同一クラスタを構成する他方の電源系統に属するチャンネル制御部110のボードへの電源供給は継続されるため、当該チャンネル制御部110に処理を引き継ぐことができる（フェイルオーバー）。

【0056】

図1におけるディスク制御部140や共有メモリ120等の、記憶デバイス制御装置100を構成する他の装置は、図6における記憶デバイス制御装置100の背面側等に装着されている。筐体の最下段に備えられたバッテリー104aは、このような共有メモリ120等へのバックアップ電源としての機能を有する。

【0057】

記憶デバイス制御装置100には、チャンネル制御部110のボード等から発生する熱を放出するためのファン（送風部）170が設けられている。ファン170は記憶デバイス制御装置100の上面部に設けられる他、チャンネル制御部110用スロットの上部にも設けられている。本実施の形態におけるファンの回転数は、毎分3600回転である。これにより、毎分およそ2600回転の回転数を有するファンが備えられた従来の記憶デバイス制御装置の場合に比べて、冷却効率が向上する。

尚、記憶デバイス制御装置100の筐体101の詳細は後述する。

【0058】

<<<チャンネル制御部（CHN）>>>

本実施の形態におけるストレージシステム600は、前述の通り、CHN110により情報処理装置1乃至3（200）からのファイルアクセス要求を受け付け、NASとしてのサービスを、情報処理装置1乃至3（200）に提供する。

【0059】

図7において、CHN110のハードウェア構成を示す。図7において、CHN110の主たるハードウェアが、図6における筐体101の側面に平行な面に投影されて模式的に示されている。この図に示すように、CHN110のハードウェアは1つのユニットで構成される。以後、このユニットのことをNASボードとも記す。本実施の形態におけるNASボードは、互いに平行で図7における上方側で重なり合った2枚の回路基板118a、118bからなる一体的にユニット化されたボード118によって構成されている。各回路基板118a、118bは、相互に接続されて1つのユニットとして構成され、これが記憶デバイス制御装置100に対して一体的に装着できる。本実施の形態においては、複数の回路基板118a、118bからなるこのようなユニットも、同一のボード118の概念に含まれる。

つまり、本実施の形態においては、回路基板118a、118bと、当該回路基板が一体的にユニット化されたボード（回路基板部）118とは、同等の意味を有する（即ち、回路基板と回路基板部とが同等の意味を有する）。従って、NASボードは、例えば、回路基板118a、118bそのものであってもよいし、当該回路基板が一体的にユニット化されたボード（回路基板部）118であってもよい。

【0060】

より具体的には、CHN110は、後述するネットワークインタフェース部111を含み図7において回路基板118aの下方側に形成されているファイルアクセス処理部111aと、後述する入出力制御部114に含まれ図7において回路基板118aの上方側に位置する回路基板118bに形成されたI/O（Input/Output）プロセッサ119とを、主として備えている。

尚、チャンネル制御部（CHN）110については、記憶デバイス制御装置100の筐体101とともに、その詳細は後述する。

【0061】

<<<チャンネル制御部（CHF及びCHA）>>>

図8において、CHF110及びCHA110のハードウェア構成を示す。C

HF110やCHA110のボードも、CHN110と同様に、それぞれ一体的にユニット化されたボードとして形成されている。CHN110のボードと同様に、前記ボード118は、一体形成された複数枚数の回路基板から構成されているようにすることもできる。またCHF110のボード及びCHA110のボードは、CHN110のボードとサイズやボード接続用コネクタ116の位置、ボード接続用コネクタ116のピン配列等に互換性をもたせている。

【0062】

CHF110及びCHA110は、ネットワークインタフェース部111、メモリ113、入出力制御部114、I/Oプロセッサ119、及びNVRAM115を備え、これらが形成された、それぞれのボードに、ボード接続用コネクタ116及び通信コネクタ117を備えている。

【0063】

ネットワークインタフェース部111は、情報処理装置200との間で通信を行なうための通信インタフェースを備えている。CHF110の場合は、例えばファイバチャネルプロトコルに従って情報処理装置200から送信されたブロックアクセス要求を受信する。CHA110の場合は、例えば、FICON（登録商標）やESCON（登録商標）、ACONARC（登録商標）、FIBARC（登録商標）のプロトコルに従って情報処理装置200から送信されたブロックアクセス要求を受信する。通信コネクタ117は、情報処理装置200と通信を行なうためのコネクタである。CHF110の場合は、SAN500に接続可能なコネクタであり、例えば、ファイバチャネルに対応している。CHA110の場合は、情報処理装置5と接続可能なコネクタであり、FICON（登録商標）やESCON（登録商標）、ACONARC（登録商標）、FIBARC（登録商標）に対応している。

【0064】

入出力制御部114は、それぞれCHF110、CHA110の全体の制御を司るとともに、ディスク制御部140やキャッシュメモリ130、共有メモリ120、管理端末160との間でデータやコマンドの授受を行なう。メモリ113に格納された各種プログラムを実行することにより本実施の形態におけるCHF

110及びCHA110の機能が実現される。入出力制御部114は、I/Oプロセッサ119やNVRAM115を備えている。I/Oプロセッサ119は、前記データやコマンドの授受を制御する。NVRAM115は、I/Oプロセッサ119の制御を司るプログラムを格納する不揮発性メモリである。NVRAM115に記憶されるプログラムの内容は、管理端末160や、後述するNASマネージャ706からの指示により書き込みや書き換えを行なうことができる。

【0065】

<<<ディスク制御部>>>

図10において、ディスク制御部140のハードウェア構成を示す。

ディスク制御部140は、一体的にユニット化されたボードとして形成されている。ディスク制御部140のボードには、インタフェース部141、メモリ143、CPU142、及びNVRAM144が形成され、さらにボード接続用コネクタ145が備えられて、これらが一体的にユニット化された回路基板として形成されている。

【0066】

インタフェース部141は、接続部150を介してチャネル制御部110等との間で通信を行なうための通信インタフェースや、記憶デバイス300との間で通信を行なうための通信インタフェースを備えている。

CPU142は、ディスク制御部140全体の制御を司るとともに、チャネル制御部140や記憶デバイス300、管理端末160との間の通信を行なう。メモリ143やNVRAM144に格納された各種プログラムを実行することにより、本実施の形態におけるディスク制御部140の機能が実現される。ディスク制御部140により実現される機能としては、記憶デバイス300の制御やRAID制御、記憶デバイス300に記憶されたデータの複製管理やバックアップ制御、リモートコピー制御等が挙げられる。

NVRAM144は、CPU142の制御を司るプログラムを格納する不揮発性メモリである。NVRAM144に記憶されるプログラムの内容は、管理端末160や、NASマネージャ706からの指示により書き込みや書き換えを行なうことができる。

また、前述のボード接続用コネクタ145が記憶デバイス制御装置100側のコネクタと嵌合されることにより、ディスク制御部140は、記憶デバイス制御装置100と電氣的に接続される。

【0067】

<<<ソフトウェア構成>>>

図11において、本実施の形態におけるストレージシステム600におけるソフトウェア構成を示す。オペレーティングシステム701上では、RAIDマネージャ708、ボリュームマネージャ707、SVPマネージャ709、ファイルシステムプログラム703、ネットワーク制御部702、バックアップ管理プログラム710、障害管理プログラム705、NASマネージャ706等のソフトウェアが動作する。

【0068】

オペレーティングシステム701上で動作するRAIDマネージャ708は、RAID制御部740に対するパラメータの設定やRAID制御部740を制御する機能を提供する。RAIDマネージャ708は、オペレーティングシステム701やオペレーティングシステム701上で動作する他のアプリケーション、もしくはSVPからパラメータや制御指示情報を受け付けて、受け付けたパラメータのRAID制御部740への設定や、RAID制御部指示情報に対応する制御コマンドの送信を行なう。

【0069】

ここで設定されるパラメータとしては、例えば、RAIDグループを構成する記憶デバイス（物理ディスク）を定義（RAIDグループの構成情報、ストライプサイズの指定等）するためのパラメータ、RAIDレベル（例えば0, 1, 5）を設定するためのパラメータ等がある。また、RAIDマネージャ708がRAID制御部740に送信する制御コマンドとしては、RAIDの構成・削除・容量変更を指示するコマンド、各RAIDグループの構成情報を要求するコマンド等がある。

【0070】

ボリュームマネージャ707は、RAID制御部740によって提供されるL

Uをさらに仮想化した仮想化論理ボリュームを、ファイルシステムプログラム703に提供する。1つの仮想化論理ボリュームは、1以上の論理ボリュームによって構成される。

【0071】

ファイルシステムプログラム703の主な機能は、ネットワーク制御部702が受信したファイルアクセス要求に指定されているファイル名とそのファイル名が格納されている仮想化論理ボリューム上のアドレスとの対応づけを管理することである。例えば、ファイルシステムプログラム703は、ファイルアクセス要求に指定されているファイル名に対応する仮想化論理ボリューム上のアドレスを特定する。

【0072】

ネットワーク制御部702は、NFS (Network File System) 711とSamba 712との2つのファイルシステムプロトコルを含んで構成される。NFS 711は、NFS 711が動作するUNIX (登録商標) 系の情報処理装置200からのファイルアクセス要求を受け付ける。一方、Samba 712はCIFS (Common Interface File System) 713が動作するWindows (登録商標) 系の情報処理装置200からのファイルアクセス要求を受け付ける。

【0073】

NASマネージャ706は、ストレージシステム600について、その動作状態の確認、設定や制御等を行なうためのプログラムである。NASマネージャ706は、Webサーバとしての機能も有し、情報処理装置200からストレージシステム600の設定や制御を行なうための設定Webページを、情報処理装置200に提供する。設定Webページは、チャネル制御部1乃至4 (110)の個々において動作するNASマネージャ706により提供される。NASマネージャ706は、情報処理装置1乃至3 (200)からのHTTP (HyperText Transport Protocol) リクエストに応じて、設定Webページのデータを情報処理装置1乃至3 (200)に送信する。情報処理装置1乃至3 (200)に表示された設定Webページを利用してシステムアドミニストレータ等によりストレージシステム600の設定や制御の指示が行われる。

【0074】

NASマネージャ706は、設定Webページに対する操作に起因して情報処理装置200から送信される設定や制御に関するデータを受信してそのデータに対応する設定や制御を実行する。これにより、情報処理装置1乃至3(200)からストレージシステム600の様々な設定や制御を行なうことができる。また、NASマネージャ706は、設定Webページの設定内容を、チャンネル制御部110上で動作するOSやアプリケーションプログラム、ディスク制御部140等に通知する。設定Webページで設定された内容は、共有LU310に管理されることもある。

【0075】

NASマネージャ706の設定Webページを利用して行なうことができる内容としては、例えば、LUの管理や設定(容量管理や容量拡張・縮小、ユーザ割り当て等)、上述の複製管理やリモートコピー(レプリケーション)等の機能に関する設定や制御(複製元のLUと複製先のLUの設定等)、後述のバックアップ管理プログラム710についての設定や制御、冗長構成されたCHNやCHF、CHA間でのクラスタの管理(フェイルオーバーさせる相手の対応関係の設定、フェイルオーバー方法等)、OSやOS上で動作するアプリケーションプログラムのバージョン管理、ウイルス検知プログラムやウイルス駆除等のデータの安全性に関する機能を提供するセキュリティ管理プログラム716の動作状態の管理や設定等がある。

【0076】

バックアップ管理プログラム710は、記憶デバイス300に記憶されているデータをLAN経由またはSAN経由でバックアップするためのプログラムである。バックアップ管理プログラム710はNDMP(Network Data Management Protocol)の機能を提供し、情報処理装置200で動作するNDMP対応のバックアップソフトウェアとLAN400を通じてNDMPに従った通信を行なう。バックアップデバイス910が情報処理装置200にSCSI経由等で接続されている場合、バックアップされるデータは情報処理装置200に一旦取り込まれてからバックアップデバイス910に送られる。バックアップデバイス910が

LANに200接続されている場合には、バックアップされるデータを、情報処理装置200を経由せずにストレージシステム600から直接バックアップデバイス910に転送することもできる。

【0077】

障害管理プログラム705は、クラスタを構成するチャンネル制御部110間でのフェイルオーバー制御を行なうためのプログラムである。

【0078】

図12において、2枚のCHN110でクラスタ180が構成されている様子を示す図を示す。図12では、CHN1（チャンネル制御部1）110とCHN2（チャンネル制御部2）110とでクラスタ180が構成されている場合を示す。

【0079】

前述したように、フェイルオーバー処理はクラスタ180を構成するチャンネル制御部110間で行われる。つまり、例えばCHN1（110）に何らかの障害が発生し処理を継続することができなくなった場合には、CHN1（110）がそれまで行っていた処理はCHN2（110）に引き継がれる。フェイルオーバー処理は、CHN1（110）とCHN2（110）により実行される障害管理プログラム705により実行される。

【0080】

CHN1（110）及びCHN2（110）は、ともに障害管理プログラムを実行し、例えば共有メモリ120に対して自己の処理が正常に行われていることを書き込むようにする。そして、相手側の前記書き込みの有無を相互に確認するようにする。相手側による書き込みが検出できない場合には、相手側に何らかの障害が発生したと判断し、フェイルオーバー処理を実行する。フェイルオーバー実行時の処理の引き継ぎは、例えば共有LU310を介して行われる。

【0081】

フェイルオーバーはこのように自動的に行われることもあるが、オペレータが管理端末を操作して手動で行われることもある。またユーザがNASマネージャ706が提供する設定Webページを利用して情報処理装置200から手動で行われることもある。フェイルオーバーを手動で行なう目的としては、耐用年数の経過

やバージョンアップ、定期診断などのためにチャネル制御部110のハードウェア（例えばNASボード）を交換する必要がある場合などがある。

【0082】

SVPMネージャ709は、管理端末160からの要求に応じて各種のサービスを管理端末160に提供する。例えば、LUの設定内容やRAIDの設定内容等のストレージシステム600に関する各種設定内容の管理端末160への提供や、管理端末160から入力されたストレージシステム600に関する各種設定の反映等を行なう。

【0083】

セキュリティ管理プログラム716は、コンピュータウイルスの検知、侵入監視、コンピュータウイルス検知プログラムの更新管理、感染したコンピュータウイルスの駆除、ファイアウォール機能等を実現する。

【0084】

===筐体、及びNASボードの詳細===

<<<筐体>>>

図6において、筐体101の第1段目ボックス102には、CHN110に給電するための電源（電源装置）104が設けられている。この電源104は、CHN110が備える後述のメモリ制御部及びCPU（後述のファイルアクセス処理部111aはこの2つを備える）に対して、12Vの直流電圧を有する電力を供給する。12Vの直流電圧を有する電力は、筐体101の正面側（外方側）から、図6の矢印abにて示されているケーブル（電力用ケーブル）108と電源コネクタ（電力用コネクタ）107とを介して供給される。

【0085】

一方、CHN110が備える後述のI/Oプロセッサ119等に対しては、電源104から、3.3V及び5Vの直流電圧を有する電力が供給される。この電力は、図6において、第1、2段目ボックス102、103の背面側にそれぞれ設けられたプラッタ（支持部）に設けられたコネクタと、CHN110のボードが有するボード接続用コネクタ（内方側コネクタ）116とを介して供給される。

【0086】

両ボックス102、103のプラッタは、エポキシ樹脂からなる層と銅箔からなる層とによる積層構造をなしている。ここで、銅箔からなる層の少なくとも一部が、前記の3.3V及び5Vに対応する電力を電源104からコネクタまで供給するための、図7に示した給電ライン116bの役割を担う。また、前記2つのプラッタの間には、銅製の板状部材からなるバス・バー（Bus Bar）が設けられている。これらの給電ラインとバス・バーとによって、電源104からCHN110にかけて、導電部が形成される。この導電部によって、電源104からCHN110へ、電流の損失をできるだけ抑制しつつ、給電が行なわれる。

【0087】

ところで、筐体101に収納されて構成される記憶デバイス制御装置100及び記憶デバイス300としては、例えばSAN対応として製品化されている従来の構成の装置を利用することができる。特に前記のCHN110のボードのコネクタ形状を従来構成の筐体に設けられているスロットにそのまま装着できる形状とすることで、従来構成の装置の利用をより簡単に行なうことができる。即ち、本実施の形態におけるストレージシステム600は、既存の製品を利用することで容易に構築することができる。

【0088】

<<<NASボード>>>

図7におけるファイルアクセス処理部111aは、図1におけるLAN400を介して情報処理装置200からファイルアクセス要求を受け付ける機能を主として有する。このような機能を実現するために、本実施の形態におけるファイルアクセス処理部111aは、ネットワークインタフェース部111、メモリ制御部113a、CPU112、及びメモリ113からなっている。

【0089】

図7に示されるように、ネットワークインタフェース部111は、通信コネクタ（通信用コネクタ）117とメモリ制御部113aとに接続されている。ネットワークインタフェース部111は、通信コネクタ117とメモリ制御部113aとの間のインタフェースの役割を担っている。本実施の形態におけるネットワ

ークインタフェース部111は、回路基板118aに形成され、両回路基板118a及び118bの間隙に位置している。通信コネクタ117は、図7における回路基板118aの右端側に2つ設けられている。ここで、図7における右端側とは、図6における筐体101の外方側、即ち正面側に相当する。また、通信コネクタ117は、図1における情報処理装置200に、LAN400を介して接続されている。さらに、本実施の形態における通信コネクタはイーサネット（登録商標）に対応している。

尚、本実施の形態における通信コネクタ117には、後述のI/Oプロセッサ119を主とした従来の構成、例えばSAN対応として製品化されている構成としてのボードの備えた通信用コネクタを、そのまま使用してもよい。

【0090】

メモリ制御部113aは、メモリ113とCPU112とに接続されている。メモリ制御部113aは、図7における回路基板118aの下方側で左右方向略中央に形成されている。メモリ制御部113aは、ネットワークインタフェース部111経由でメモリ113及び通信コネクタ117間のデータの転送と、メモリ113及びCPU112間のデータの転送とを制御する機能を有する。本実施の形態においては、メモリ制御部113aには、2.5Vの電圧が印加される。この電圧は、第1DC・DCコンバータ（電圧変換部）105から印加される。この第1DC・DCコンバータ105は、図7における回路基板118aの下方側で右側（図6における筐体101の外方側）に形成され、電源104から入力される12Vの電圧を2.5Vの電圧に変換して出力する。

【0091】

CPU112は、図7における回路基板118aの下方側で左右方向略中央に形成され、CHN110のボードをNASボードとして機能させるための制御を司る。本実施の形態においては、CPU112には、1.3Vの電圧が印加される。この電圧は、第2DC・DCコンバータ（電圧変換部）106から印加される。この第2DC・DCコンバータ106は、図7における回路基板118aの下方側で右側（図6における筐体101の外方側）に形成され、電源104から入力される12Vの電圧を1.3Vの電圧に変換して出力する。

【0092】

メモリ113は、図7における回路基板118aの下方で左側（図6における筐体101の内方側）に2つ形成されている。メモリ113には、様々なプログラムやデータが記憶される。例えば、図9に示すメタデータ730やロックテーブル720、また図11に示すNASマネージャ706等の各種プログラムが記憶される。メタデータ730は、ファイルシステムが管理しているファイルに対応させて生成される情報である。メタデータ730には、例えば、ファイルのデータが記憶されているLU上のアドレスやデータサイズ等、ファイルの保管場所を特定するための情報が含まれる。メタデータ730には、ファイルの容量、所有者、更新時刻等の情報が含まれることもある。また、メタデータ730はファイルだけでなくディレクトリに対応させて生成されることもある。

図13において、メタデータ730の例を示す。メタデータ730は、記憶デバイス300上の各LUにも記憶されている。

【0093】

ロックテーブル720は、情報処理装置1乃至3（200）からのファイルアクセスに対して排他制御を行なうためのテーブルである。排他制御を行なうことにより情報処理装置1乃至3（200）でファイルを共用することができる。

図14において、ロックテーブル720を示す。図14に示すように、ロックテーブル720には、ファイルロックテーブル721とLUロックテーブル722とがある。ファイルロックテーブル721は、ファイル毎にロックが掛けられているか否かを示すためのテーブルである。いずれかの情報処理装置200により、あるファイルがオープンされている場合に、当該ファイルにロックが掛けられる。ロックが掛けられたファイルに対する他の情報処理装置200によるアクセスは禁止される。LUロックテーブル722は、LU毎にロックが掛けられているか否かを示すためのテーブルである。いずれかの情報処理装置200により、あるLUに対するアクセスが行われている場合に、当該LUにロックが掛けられる。ロックが掛けられたLUに対する他の情報処理装置200によるアクセスは禁止される。

【0094】

本実施の形態においては、メモリ113には、第1DC・DCコンバータ105から、2.5Vの電圧が印加される。

【0095】

一方、図7に示される入出力制御部114は、図1におけるディスク制御部140やキャッシュメモリ130、共有メモリ120、及び管理端末160との間でデータやコマンドの授受を行なう。入出力制御部114は、I/Oプロセッサ119やNVRAMを備えている。I/Oプロセッサ119は、例えば1チップのマイコンで構成される。I/Oプロセッサ119は、前記データやコマンドの授受を制御し、CPU112とディスク制御部140との間の通信を中継する。本実施の形態においては、2つのI/Oプロセッサ119が、回路基板118bに形成されている。NVRAMは、I/Oプロセッサ119の制御を司るプログラムを格納する不揮発性メモリである。NVRAMに記憶されるプログラムの内容は、図1及び6に示される管理端末160や、NASマネージャ706からの指示により書き込みや書き換えを行なうことができる。

【0096】

電源104からの前記の12Vの電圧は、図6及び7における電源コネクタ107とケーブル108とによって、CHN110まで供給される。この電源コネクタ107は、図7における回路基板118aの右側端部（筐体の外方側に位置すべき端部）に設けられている。

【0097】

また、前記I/Oプロセッサ119等には、ボード118に適宜に形成されたDC・DCコンバータから電力が供給される。そして、この適宜に形成されたDC・DCコンバータには、図7における回路基板118aの左側端部に設けられたボード接続用コネクタ（内方側コネクタ）116と、ブラッタ116a内部の給電ライン116bとを介して、電源104から3.3V及び5Vの電圧を有する電力が供給される。さらに、前記のボード接続用コネクタ116は、I/Oプロセッサ119等に対する電力の供給のみならず、当該I/Oプロセッサと図1における内部LAN151等との接続のための機能も有する。

【0098】

本実施の形態における、CPU 112に対する印加電圧は1.3Vであるが、動作中のCPU 112における電流の変動が大きい。一方、ボード接続用コネクタ 116を通じて供給される3.3V及び5Vの電圧を有する電力では、このCPU 112の電流変動に対応することが困難である。よって、この大きな電流変動に対応すべく、12Vの電圧を有する電力が必要となり、これが電源コネクタ 107を通じて供給される。

【0099】

前述したように、図6及び7におけるCHN 110用スロットの上部と筐体 101の上面とにはそれぞれファン 170が設けられている。これにより、図7の矢印で示される下側から上側に向けて空気流が形成され、CHN 110のハードウェア構成が冷却される。CHN 110用スロットの上部においては、ファン 170は、CHN 110のボード1つにつき2つ対応するように設けられている。

【0100】

本実施の形態における、ボード接続用コネクタ 116を有するボード 118には、I/Oプロセッサ 119を主とした従来の構成、例えばSAN対応として製品化されている構成を適用することができる。本実施の形態のCHN 110のボード（NASボード）においては、従来の構成に対してファイルアクセス処理部 111aが新たに設けられ、図7における回路基板 118aの右側端部に電源コネクタ 107が設けられている。また、ファイルアクセス処理部 111aと情報処理装置 200とを接続する通信コネクタ 117も、図7における回路基板 118aの右側端部に設けられている。よって、従来のボード接続用コネクタ 116を、新たに形成するファイルアクセス処理部 111aへの給電のために改造する必要がほとんどない。従って、このCHN 110のボード（NASボード）を、前述した従来構成の筐体に設けられているスロットにそのまま装着することによって、従来構成の装置の利用をより簡単に行なうことができる。即ち、本実施の形態における記憶デバイス制御装置 100は、既存の製品を利用することで容易に構築することができる。これにより、コンパクトで低コストな、拡張時にユーザへの負担の軽い記憶デバイス制御装置 100、及びCHN 110のボード（NASボード）を提供できる。そして、記憶デバイス制御装置 100は、コンパク

トになることによって相対的に高密度実装となる。

【0101】

例えば、従来構成による既存の筐体の有するプラッタが、I/Oプロセッサ119等に給電するための前述した銅箔の層からなる給電ラインを備えているとした場合、当該プラッタに、新たなファイルアクセス処理部111aのための同様の給電ラインを増設することは容易ではない。従って、このような増設をすることなく、既存の筐体を使用できる、本実施の形態における記憶デバイス制御装置100、及びCHN110のボード（NASボード）は、より低コストで、拡張時にユーザへの負担のより軽いものとなる。

【0102】

また、本実施の形態のボード118においては、ファイルアクセス処理部111aへの給電のための電源コネクタ107が、I/Oプロセッサ119等への給電のためのボード接続用コネクタ116に対して反対側に設けられている。これにより、ファイルアクセス処理部111aとI/Oプロセッサ119との近接配置が回避される。ここで、ファイルアクセス処理部111aの有する特にメモリ制御部113a及びCPU112と、I/Oプロセッサ119とは発熱体である。従って、本実施の形態においては、前記2つのコネクタ107、116の配置により、発熱体間の間隙が広がっている。この間隙に、ファン170によって形成される空気流が生じると、これらの発熱体に対する冷却効率が高まる。従って、記憶デバイス制御装置100の動作が安定する。

【0103】

さらに、本実施の形態においては、ファン170によって下側から上側に空気が流れている。よって、ボード118において、下側の方が上側よりも温度が低い。従って、ボード118の上側にI/Oプロセッサ119等を配置し、ボード118の下側に発熱量がより大きいメモリ制御部113a及びCPU112を配置する。これにより、CHN110のボード（NASボード）全体を効率良く冷却できるため、記憶デバイス制御装置100の動作が安定する。

【0104】

加えて、本実施の形態においては、一対の回路基板118a、118bどうし

の間隙にファン170からの送風を受けて、CHN110のボード（NASボード）全体がより効率良く冷却されるため、記憶デバイス制御装置100の動作がより安定する。

【0105】

前述したように、本実施の形態のボード118においては、ファイルアクセス処理部111aへの給電のための電源コネクタ107が、I/Oプロセッサ119等への給電のためのボード接続用コネクタ116に対して反対側に設けられている。ここで、第1、2DC・DCコンバータ105、106と、I/Oプロセッサ119等のために適宜に形成されたDC・DCコンバータとが、電力の損失を抑制するために、電源コネクタ107の近傍と、ボード接続用コネクタ116の近傍とにそれぞれ分けて配置されんとする。このような、コネクタ107、116の分散配置により、DC・DCコンバータ105、106と、前記の適宜に形成されたDC・DCコンバータとの分散配置が実現する。本実施の形態においては、図7における左右方向に関して、DC・DCコンバータ105、106は、ファイルアクセス処理部111aを構成するメモリ制御部113a及びCPU112と、電源コネクタ107との間に位置している。従って、これらのDC・DCコンバータ105、106は、I/Oプロセッサ119等のために適宜に形成されたDC・DCコンバータとは離れた位置にある。ここで、一般に、DC・DCコンバータは高周波を発生させて、通信コネクタ117やボード接続用コネクタ116における電気信号に対してノイズやグラウンドレベルのシフトを生じさせる恐れがあるとされている。従って、本実施の形態におけるDC・DCコンバータ105、106と、前記の適宜に形成されたDC・DCコンバータとの分散配置によって、前記のノイズやシフトの発生源を分散することができる。これにより、記憶デバイス制御装置100の動作が安定する。

【0106】

また、本実施の形態においては、ケーブル108という簡便な手段により、電源104から電源コネクタ107へ12Vの電圧を印加している。これにより、低コストな記憶デバイス制御装置100、及びCHN110のボード（NASボード）を提供できる。

【0107】

===その他の実施の形態===

前述した発明の実施の形態は、本発明の理解を容易にするためのものであり、本発明を限定して解釈するためのものではない。本発明は、その趣旨を逸脱することなく変更、改良され得るとともに、本発明にはその等価物も含まれる。

【0108】

前記の実施の形態においては、電力は、電源104からケーブル108を介して電源コネクタ107へ供給されたとしたが、これに限定されるものではない。例えば、電源コネクタ107を介さずに、ケーブル108が直接DC・DCコンバータ105、106等に接続されていてもよい。

【0109】

また、前記の実施の形態においては、CHN110のボード（NASボード）におけるファイルアクセス処理部111aとI/Oプロセッサ119とが、互いに平行で重なった状態にある2枚の回路基板118a、118bを一組とするボード118に、互いに分けて形成されているとしたが、これに限定されるものではない。もし、ファイルアクセス処理部111aやI/Oプロセッサ119等がより小型化されれば、これら全てが、例えば1枚の回路基板118aに形成されていてもよい。

【0110】

【発明の効果】

高密度実装、低コストであり、拡張時にユーザへの負担が軽く、安定な動作を確保できる。

【図面の簡単な説明】

【図1】

本実施の形態におけるストレージシステムの全体構成を示すブロック図である。

【図2】

本実施の形態における管理端末の構成を示すブロック図である。

【図3】

本実施の形態における物理ディスク管理テーブルを示す図である。

【図 4】

本実施の形態における LU 管理テーブルを示す図である。

【図 5】

本実施の形態におけるストレージシステムの外観構成を示す図である。

【図 6】

本実施の形態における記憶デバイス制御装置の外観構成を示す図である。

【図 7】

本実施の形態における CHN を示す図である。

【図 8】

本実施の形態における CHF、CHA を示す図である。

【図 9】

本実施の形態におけるメモリに記憶されるデータの内容を説明するための図である。

【図 10】

本実施の形態におけるディスク制御部を示す図である。

【図 1.1】

本実施の形態におけるソフトウェア構成図である。

【図 1.2】

本実施の形態におけるチャネル制御部においてグラスタが構成されている様子
を示す図である。

【図 1.3】

本実施の形態におけるメタデータを示す図である。

【図 1.4】

本実施の形態におけるロックテーブルを示す図である。

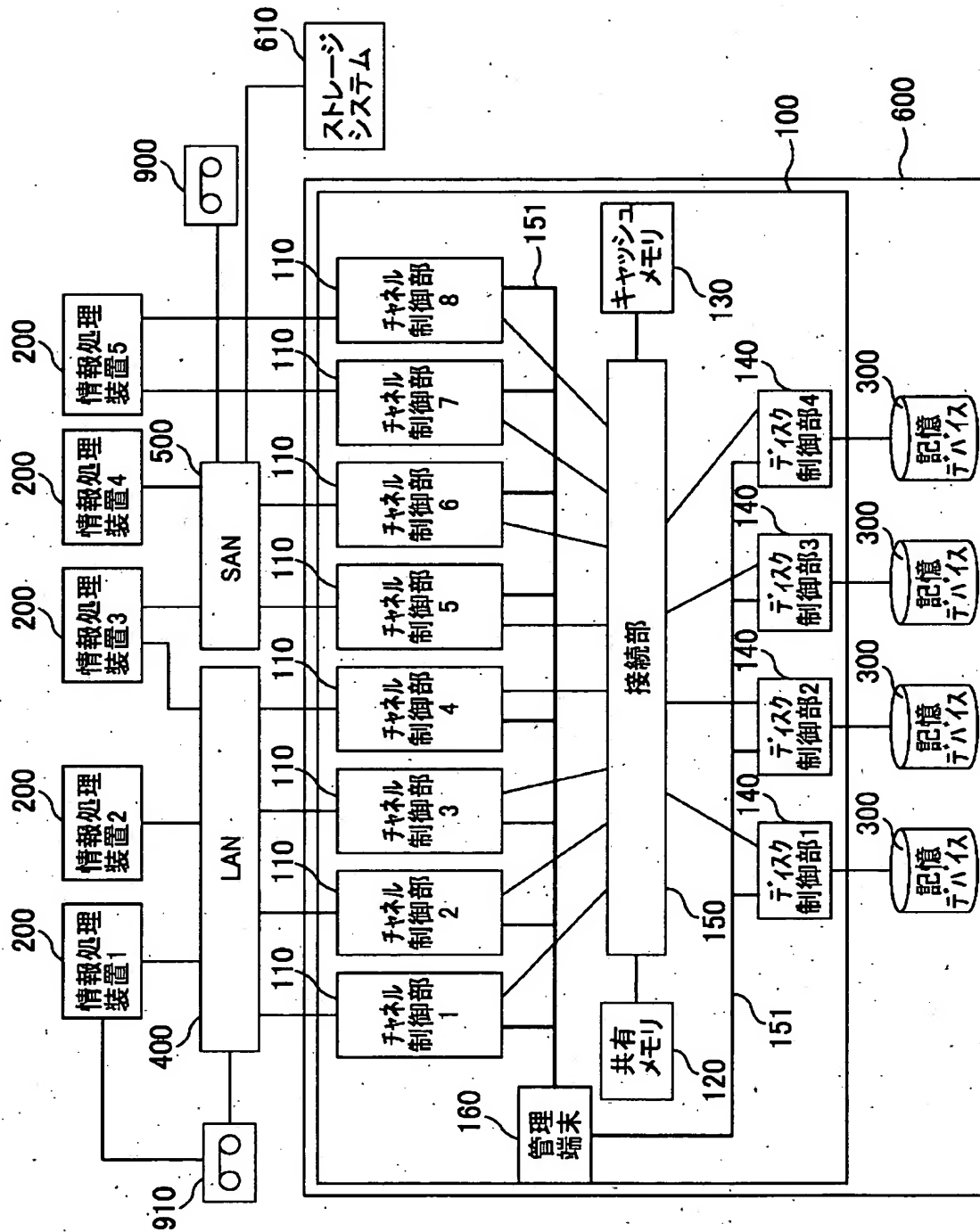
【符号の説明】

- | | | | |
|------|------------|------|----------|
| 1.00 | 記憶デバイス制御装置 | 101 | 筐体 |
| 102 | 第1段目ボックス | 103 | 第2段目ボックス |
| 104 | 電源 | 104a | バッテリー |

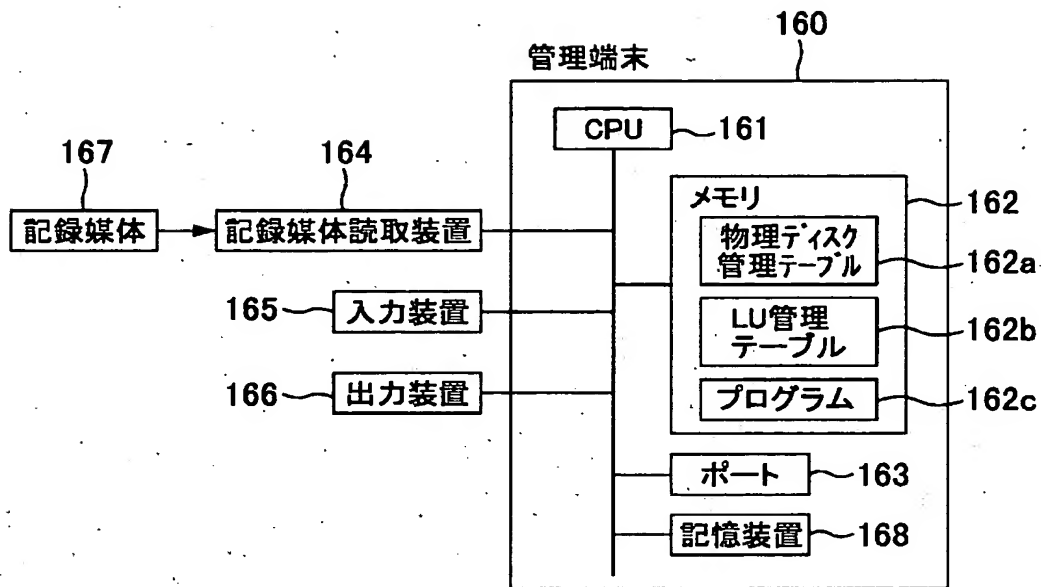
105	第1DC・DCコンバータ	106	第2DC・DCコンバータ
107	電源コネクタ	108	ケーブル
110	チャンネル制御部	111	ネットワークインタフェース部
111a	ファイルアクセス処理部	112	CPU
113	メモリ	113a	メモリ制御部
114	入出力制御部	116	ボード接続用コネクタ
116a	プラッタ	116b	給電ライン
117	通信コネクタ	118	ボード
118a、118b	回路基板	119	I/Oプロセッサ
120	共有メモリ	130	キャッシュメモリ
140	ディスク制御部	150	接続部
151	内部LAN	160	管理端末
170	ファン	180	クラスタ
200	情報処理装置	300	記憶デバイス
400	LAN	500	SAN
600	ストレージシステム	721	ファイルロックテーブル
722	LUロックテーブル	730	メタデータ
900	SAN対応バックアップデバイス		
910	バックアップデバイス (テープデバイス)		

【書類名】 図面

【図 1】



【図 2】



【図 3】

162a

物理ディスク管理テーブル

ディスク番号	容量	RAID	使用状況
#001	100GB	5	使用中
#002	100GB	5	使用中
#003	100GB	5	使用中
#004	100GB	5	使用中
#005	100GB	5	使用中
#006	50GB	—	未使用
⋮	⋮	⋮	⋮

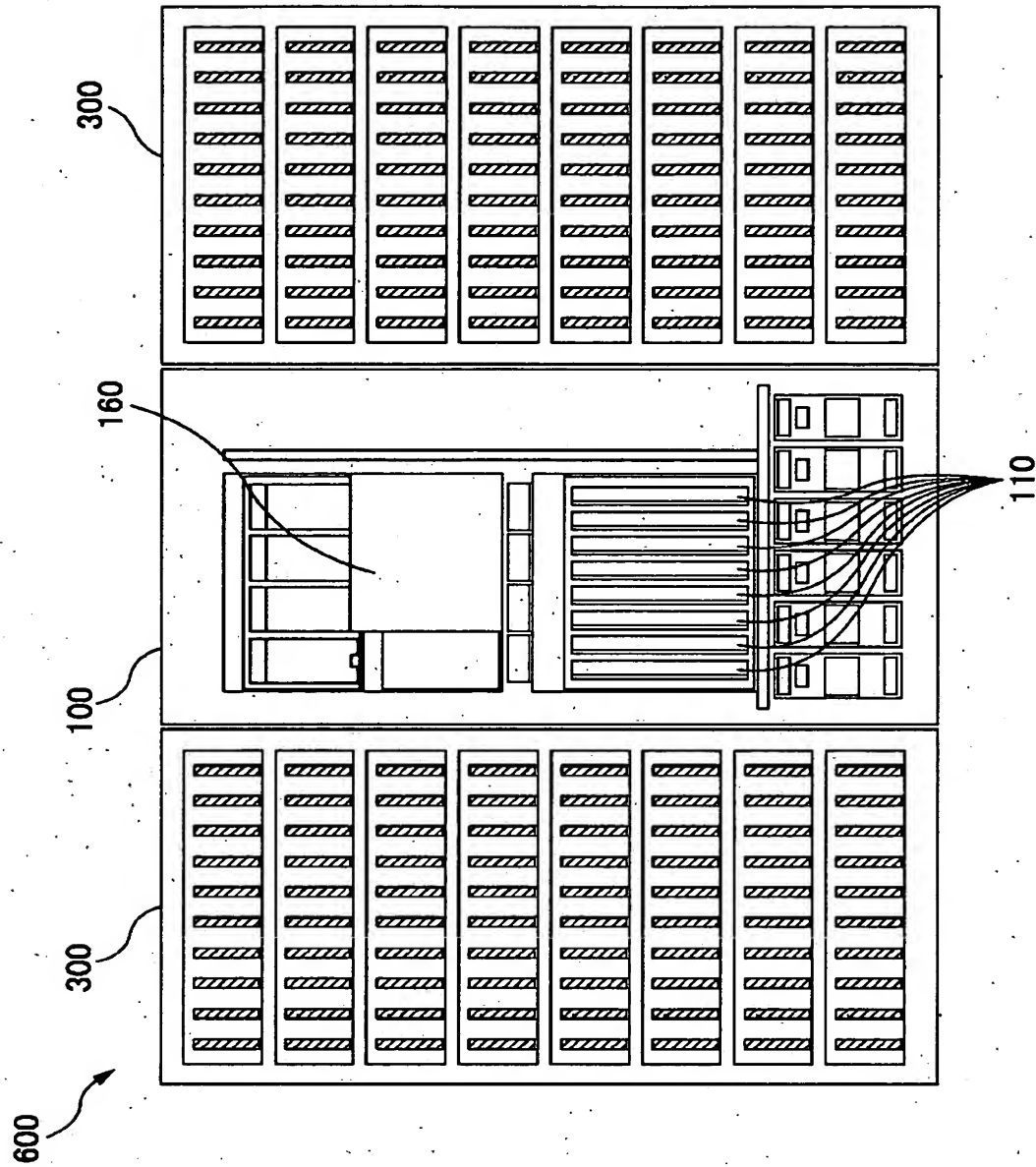
【図 4】

162b

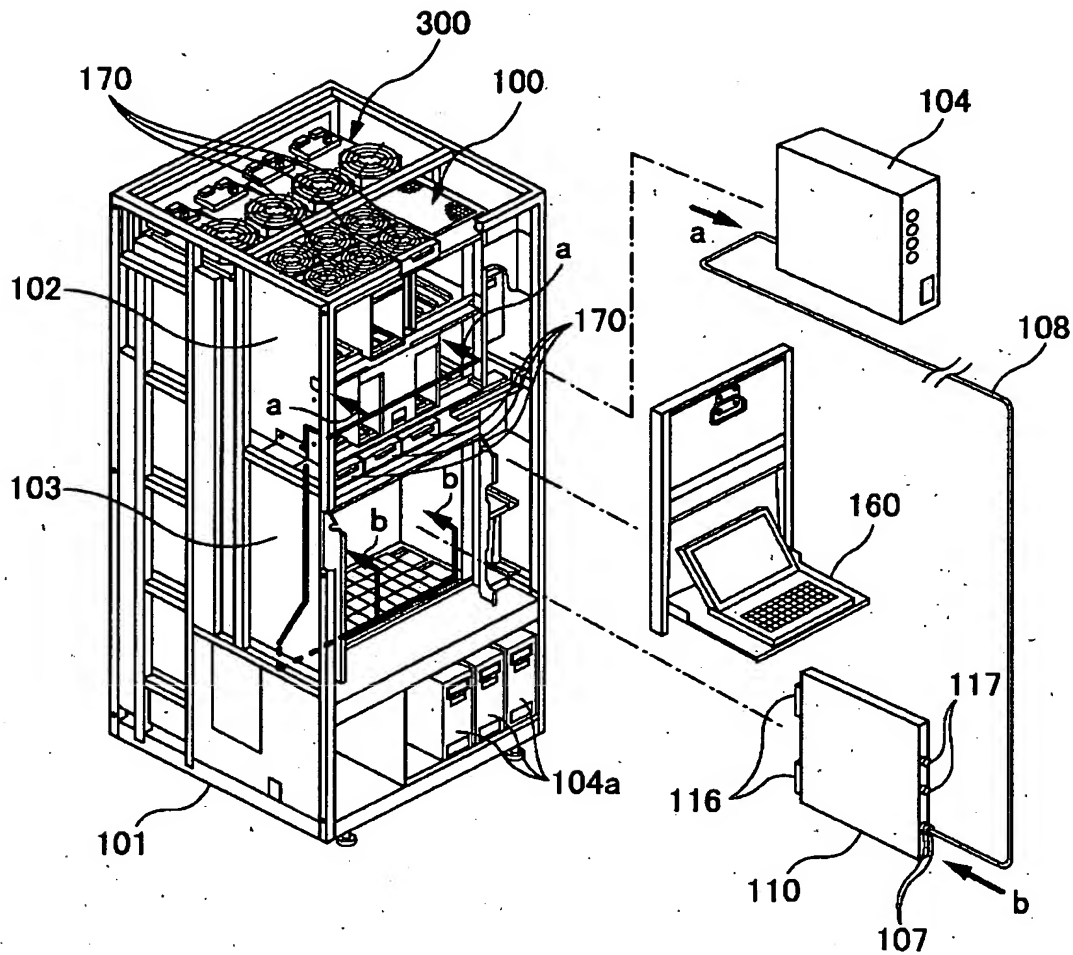
LU管理テーブル

LU番号	物理ディスク	容量	RAID
#1	#001,#002,#003,#004,#005	100GB	5
#2	#001,#002,#003,#004,#005	300GB	5
#3	#006,#007,	200GB	1
⋮	⋮	⋮	⋮

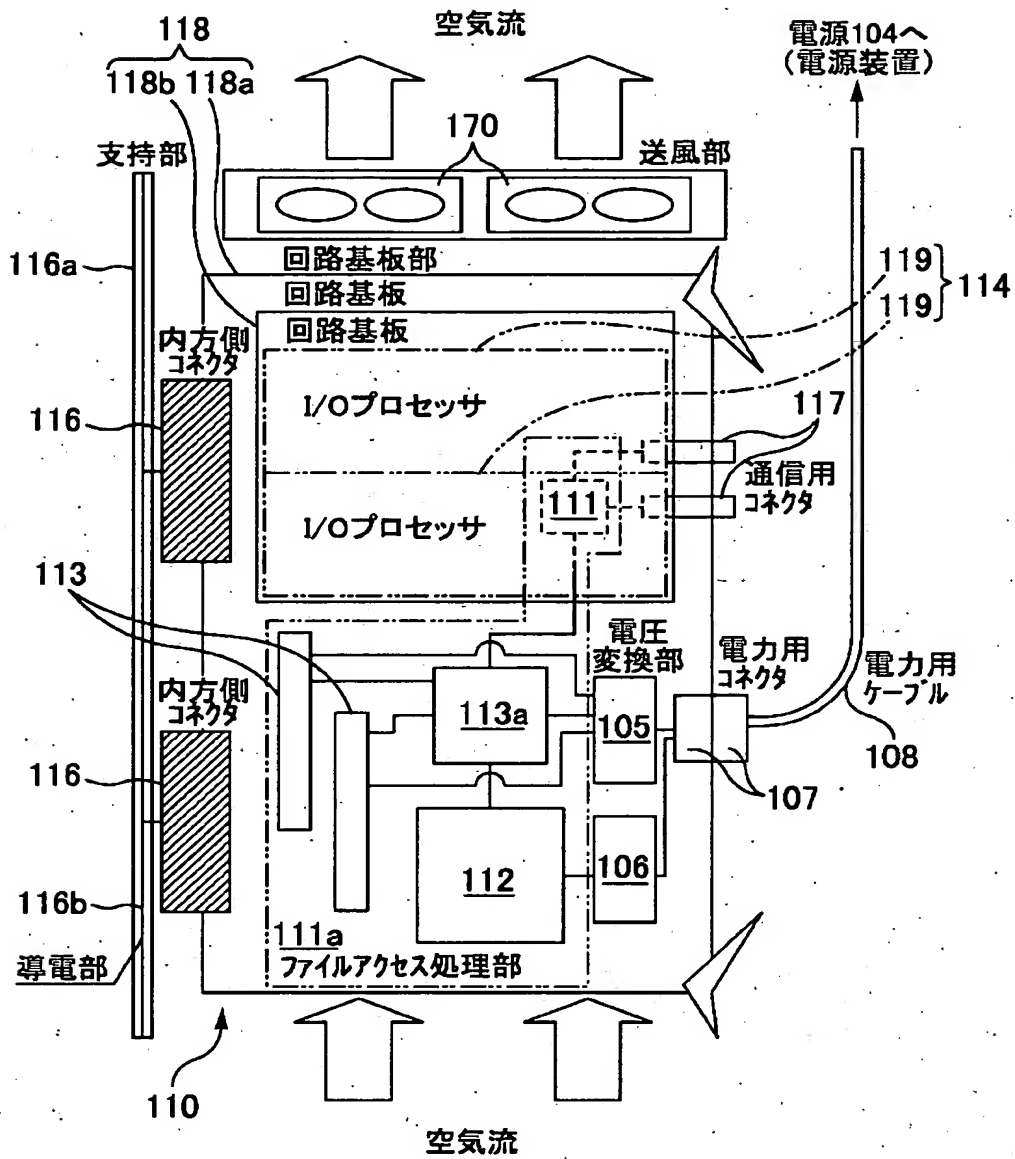
【図 5】



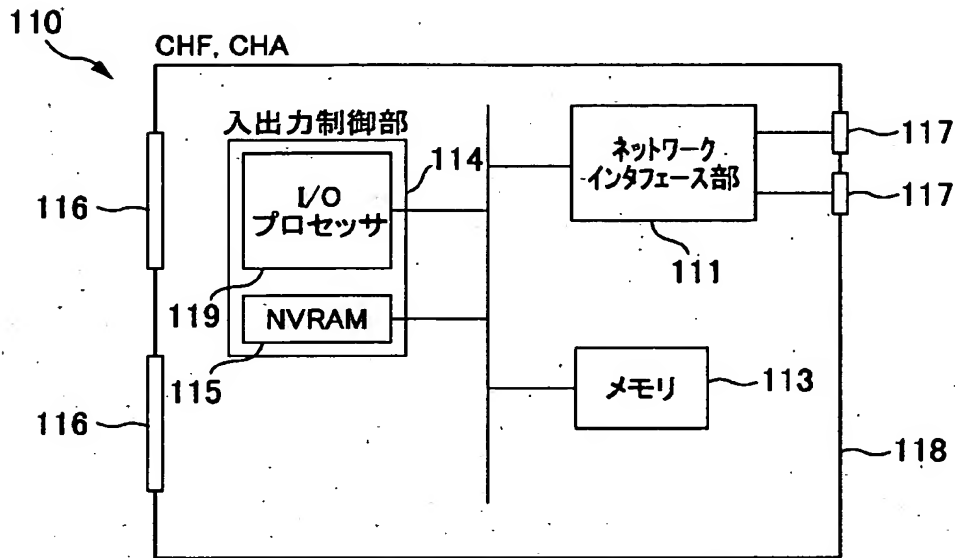
【図 6】



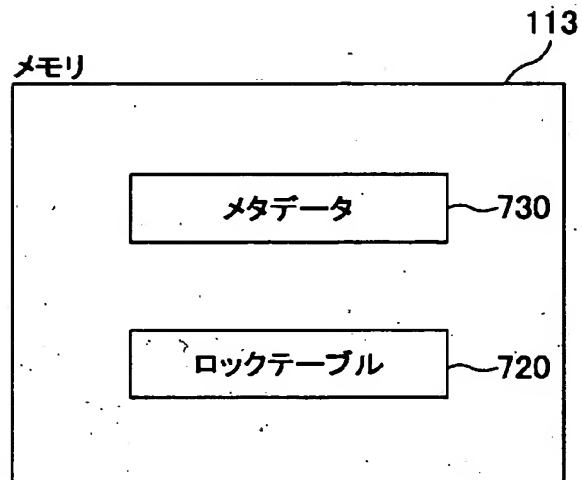
【図 7】



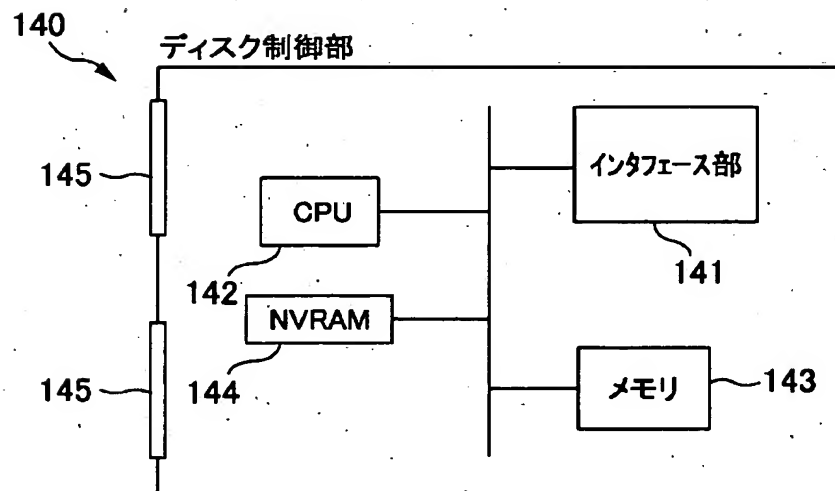
【図 8】



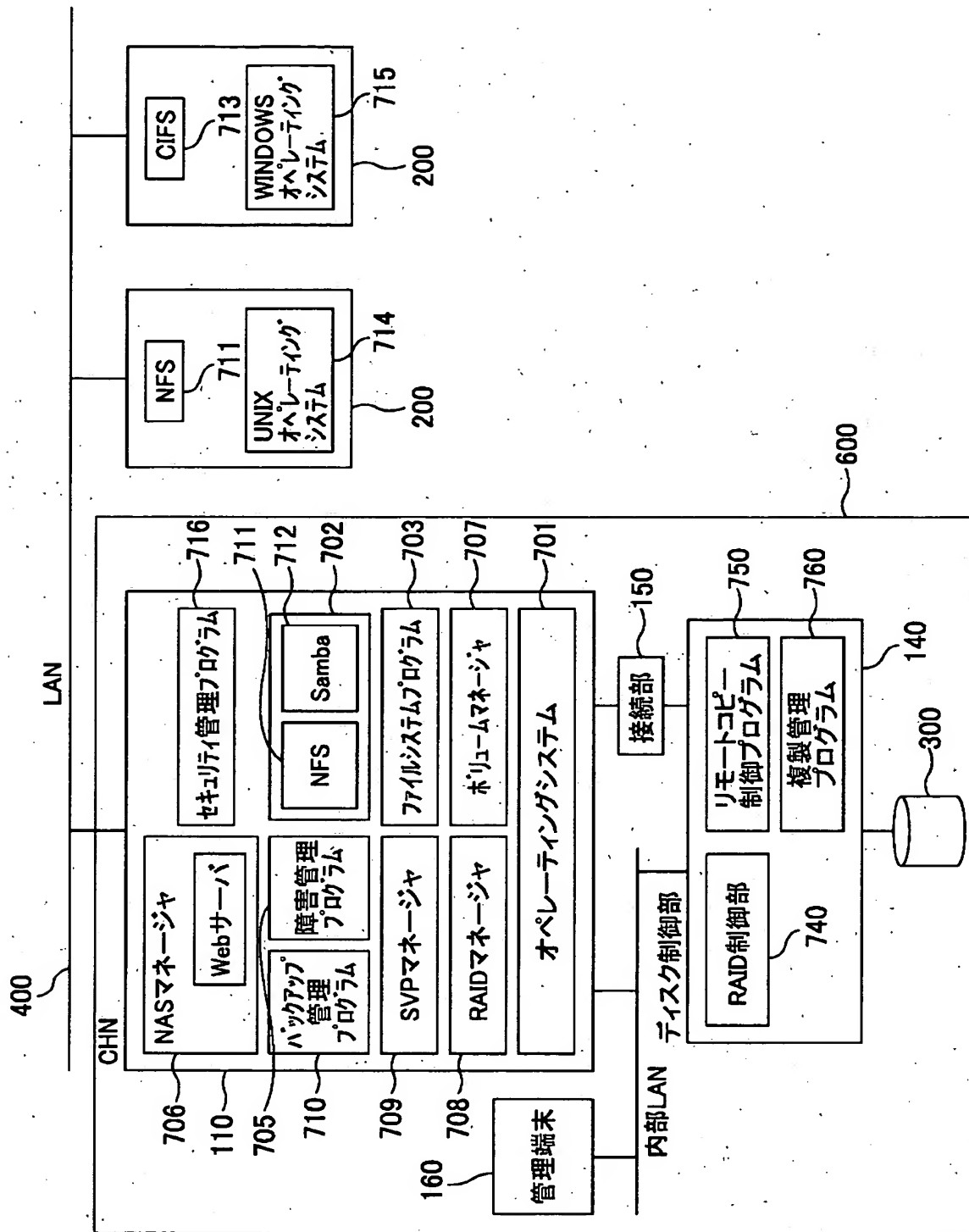
【図 9】



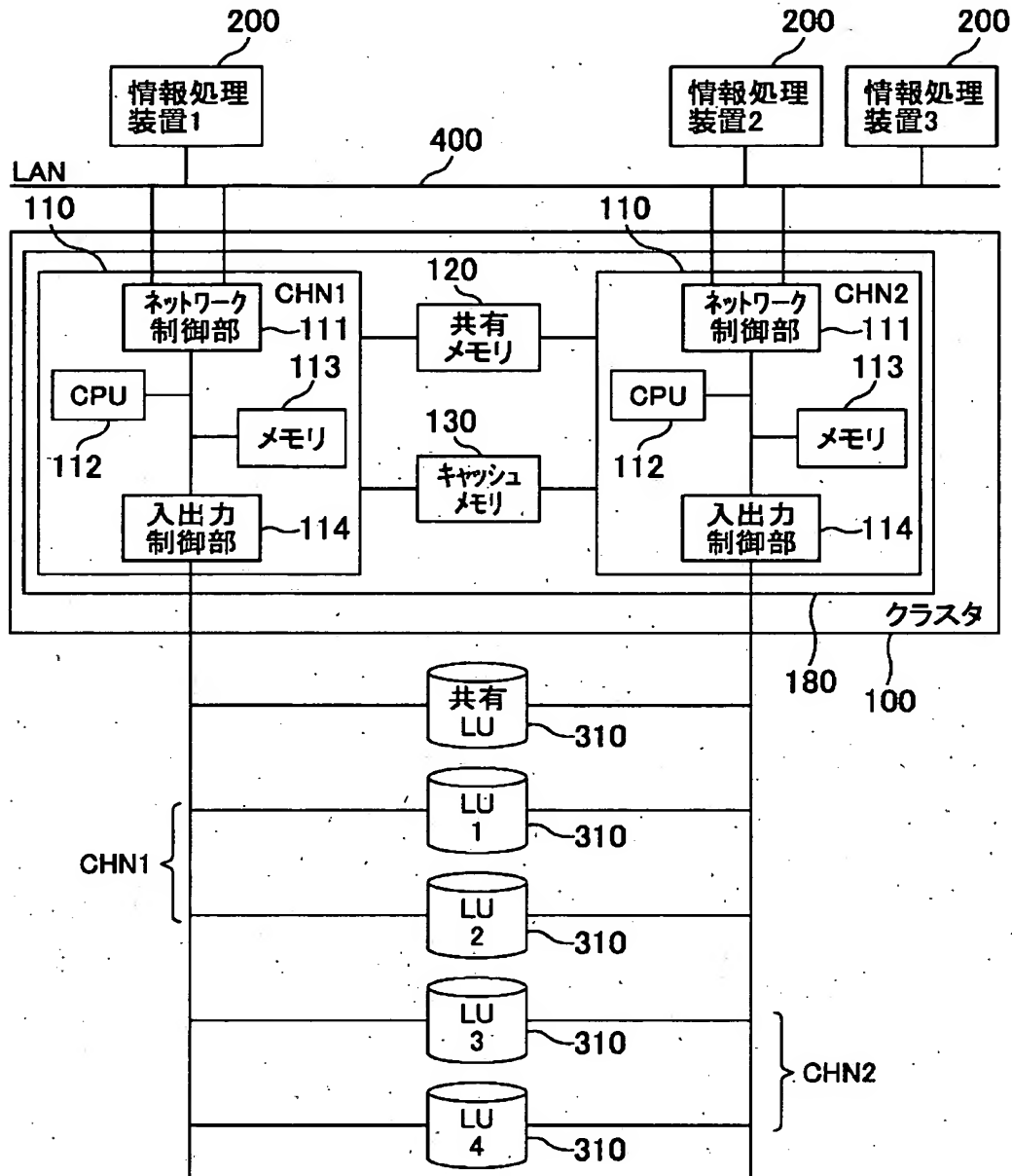
【図 10】



【図11】



【図 12】



【図 13】

730

メタデータ

ファイル名	先頭アドレス	容量	所有者	更新時刻
A	7BSA	200MB	X	0:00
B	05BF	50MB	X	7:57
C	1F30	100MB	Y	9:15
D	470B	100MB	Z	15:20
⋮	⋮	⋮	⋮	⋮

【図 14】

721

ファイルロックテーブル

ファイル名	ロック状態
A	ロック中
B	—
C	—
D	ロック中
⋮	⋮

722

LUロックテーブル

LU	ロック状態
共有	—
1	ロック中
2	—
⋮	⋮

【書類名】 要約書

【要約】

【解決手段】 データを記憶する記憶デバイスに接続されるとともに、ネットワークを介して情報処理装置と接続され、筐体に収納される記憶デバイス制御装置であって、情報処理装置からのデータ入出力要求に対応するI/O要求を、記憶デバイスに対して出力するI/Oプロセッサが形成された、筐体に収納される回路基板を有し、回路基板は、筐体の内方側に位置すべき端部に内方側コネクタを有し、内方側コネクタを通じて、少なくともI/Oプロセッサと電源装置とが接続される、記憶デバイス制御装置において、回路基板には、データ入出力要求をファイル単位で受け付けるファイルアクセス処理部が形成され、回路基板は、筐体の外方側に位置すべき端部にファイルアクセス処理部へ電力を供給するための電力用コネクタを備えている。

【選択図】 図7

出願人履歴情報

識別番号

[000005108]

1. 変更年月日	1990年 8月31日
[変更理由]	新規登録
住 所	東京都千代田区神田駿河台4丁目6番地
氏 名	株式会社日立製作所